

Chapter 2

Overlapping Singer Identification in Music Recordings

Wei-Ho Tsai

National Taipei University of Technology, Taiwan

Cin-Hao Ma

National Taipei University of Technology, Taiwan

ABSTRACT

Singer identification (SID), which refers to the task of automatically identifying the singer(s) in a music recording, is of great help in handling the rapid proliferation of music data on the internet and digital media. Although a number of SID studies from acoustic features have been reported, most systems are designed to identify the singer in recordings of solo performances. Very little research has considered a more realistic case, which is to identify more than one singer in a music recording. The research presented in this chapter investigates the feasibility of identifying singers in music recordings that contain overlapping (simultaneous) singing voices (e.g., duet or trio singings). This problem is referred to as overlapping singer identification (OSID). Several approaches to OSID are discussed and evaluated in this chapter. In addition, a related issue on how to distinguish solo singings from overlapping singing recordings is also discussed.

DOI: 10.4018/978-1-5225-7268-8.ch002

INTRODUCTION

Explosive growth in the Internet and digital media has motivated recent research on developing techniques for automatically extracting information from music for content-based retrieval (Michael *et al.*, 2008; Schedl *et al.*, 2014). In music recordings, the singing voice usually catches more of listeners' attention than other music attributes such as rhythm, tonality, or instrumentation. Therefore, extracting information on singers is essential to people for organizing, browsing, and retrieving music recordings, especially for singer identity information where it may be undocumented or difficult to find, such as cameo's or guest appearances in live concert recordings or a movie's musical interludes. In addition, singer information could enable rapidly scanning suspect websites for piracy, especially for bootleg concert recordings, in which the company will typically not have a copy of the original audio data for comparison.

Most people use singer's voice as a primary cue for identifying songs, and performing such a task is almost effortless. However, building a practicable automatic singer identification (SID) system (Tsai & Lin, 2011) is not an easy task for machine learning. One of the challenges lies in training the system to discriminate among the different sources of sounds intertwined in music recordings, which may include background vocal, instrumental accompaniment, background noise, and overlapping singings.

In the earlier period, a number of studies on SID focused on exploiting various musical features and combining them for SID (Kim & Whitman, 2002; Berenzweig *et al.*, 2002; Liu & Huang, 2002; Zhang, 2003; Bartsch & Wakefield, 2004; Tsai *et al.*, 2004; Maddage *et al.*, 2004; Fujihara *et al.*, 2005; Mesaros & Astola, 2005; Nwe & Li, 2007; Mesaros *et al.*, 2007; Nwe & Li, 2008). More recently, SID research has shifted the attentions to investigating the influence of background sounds on singer voice characterization, and several methods have been developed to reduce the interference of background sounds for SID (Tsai & Wang, 2006; Fujihara *et al.*, 2010; Tsai & Lin, 2011; Hu & Liu, 2015). However, very little research has considered the problem of automatically identifying more than one singer in a music recording.

Tsai & Wang (2004) investigated automatic detection and tracking of multiple singers in music recordings. However, the study only considered singing by multiple singers who performed in non-overlapping matters, that is, did not consider multiple voices singing simultaneously. By contrast, Tsai & Ma (2017) proposes a system to automatically identify multiple singers in a long audio stream that may have singing voices overlapping in time. The study explicitly discussed the problem of multiple voices singing simultaneously. In line with the research goal of Tsai & Ma (2017), the research presented in this chapter focuses on the problem of automatically identifying singers in music recordings that contain both simultaneous and non-simultaneous singings. We refer to this problem as overlapping singer identification (OSID). Other

19 more pages are available in the full version of this document, which may be purchased using the "Add to Cart" button on the publisher's webpage: www.igi-global.com/chapter/overlapping-singer-identification-in-music-recordings/214830

Related Content

XML Data Services

Vinayak Borkar, Michael Carey, Nitin Mangtani, Denny McKinney, Rahul Patel and Sachin Thatte (2006). *International Journal of Web Services Research* (pp. 85-95). www.irma-international.org/article/xml-data-services/3076

Reputation Management for Composite Services in Service-Oriented Systems

Surya Nepal, Zaki Malik and Athman Bouguettaya (2011). *International Journal of Web Services Research* (pp. 29-52). www.irma-international.org/article/reputation-management-composite-services-service/55235

Constructing Home Network Systems and Integrated Services Using Legacy Home Appliances and Web Services

Masahide Nakamura, Akihiro Tanaka, Hiroshi Igaki, Haruaki Tamada and Ken-ichi Matsumoto (2008). *International Journal of Web Services Research* (pp. 82-98). www.irma-international.org/article/constructing-home-network-systems-integrated/3116

Using Geospatial Web Services Holistically in Emergency Management

Ning An, Gang Liu and Baris Kazar (2011). *Geospatial Web Services: Advances in Information Interoperability* (pp. 401-425). www.irma-international.org/chapter/using-geospatial-web-services-holistically/51496

Survey Paper on Semantic Web

Rimpal Unadkat (2019). *Web Services: Concepts, Methodologies, Tools, and Applications* (pp. 2138-2143). www.irma-international.org/chapter/survey-paper-on-semantic-web/217934